Journal of the Korean Statistical Society (2002), **31**: *3, pp 315–328*

Minimax Choice and Convex Combinations of Generalized Pickands Estimator of the Extreme Value Index

Seokhoon Yun[1]

ABSTRACT

As an extension of the well-known Pickands (1975) estimator for the extreme value index, Yun (2002) introduced a generalized Pickands estimator. This paper searches for a minimax estimator in the sense of minimizing the maximum asymptotic relative efficiency of the Pickands estimator with respect to the generalized one. To reduce the asymptotic variance of the resulting estimator, convex combinations of the minimax estimator are also considered and their asymptotic normality is established. Finally, the optimal combination is determined and proves to be superior to the generalized Pickands estimator.

Keywords. Extreme value index, generalized Pickands estimator, asymptotic normality, asymptotic relative efficiency.

AMS 2000 subject classifications. Primary 62G32; Secondary 62G30.

1. Introducton

Suppose we are given *iid* observations $X_1, ..., X_n$ whose common distribution function F belongs to the domain of attraction of an extreme value distribution G_β for some $\beta \in \mathbb{R}$ [$F \in \mathcal{D}(G_\beta)$], where $G_\beta(x) := \exp\{-(1+\beta x)^{-1/\beta}\}$, $1+\beta x > 0$. This means that there exist constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G_\beta(x) \tag{1.1}$$

for all x with $1 + \beta x > 0$. Throughout the case $\beta = 0$ is interpreted as the limit when $\beta \to 0$, so that $G_0(x) = \exp(-e^{-x})$, $x \in \mathbb{R}$.

The shape parameter β of the extreme value distribution G_β, which is called the extreme value index, can be estimated using a number of upper order statistics

Received February 2002; accepted March 2002.

[1]Department of Applied Statistics, University of Suwon, Suwon, Kyonggi-do 445-743, Korea

of the sample $X_1, ..., X_n$. If it is known that $\beta > 0$, the most frequently used estimator of β is the Hill (1975) estimator, which cannot however be generally applicable for estimating general $\beta \in \mathbb{R}$ due to its inconsistency for $\beta < 0$. A popular estimator of general $\beta \in \mathbb{R}$ is the Pickands (1975) estimator, which, based on the m $(4 \leq m \leq n)$ upper order statistics, is defined by

$$\hat{\beta}_{n,m}^{(P)} := \frac{1}{\log 2} \log \frac{X_{[m/4]}^{(n)} - X_{[m/2]}^{(n)}}{X_{[m/2]}^{(n)} - X_m^{(n)}},$$

where $X_1^{(n)} \geq X_2^{(n)} \geq \cdots \geq X_n^{(n)}$ are the descending order statistics of $X_1, ..., X_n$ and $[x]$ denotes the integer part of $x \in \mathbb{R}$.

Falk (1994) and Drees (1995) considered convex combinations of several Pickands estimators to reduce the asymptotic variance of $\hat{\beta}_{n,m}^{(P)}$. On the other hand, a number of attempts to generalize the Pickands estimator have been also made in the literature (*e.g.* see Pereira, 1994; Alves, 1995; Yun, 2000). The most general form was introduced by Yun (2002), who considered the generalized Pickands estimator of form

$$\hat{\beta}_{n,m}(u,v) := \frac{1}{\log v} \log \frac{X_m^{(n)} - X_{[um]}^{(n)}}{X_{[vm]}^{(n)} - X_{[uvm]}^{(n)}}, \ u, v \in (0,1),$$

where $1 \leq m, [um], [vm], [uvm] \leq n$. The Pickands estimator corresponds to $\hat{\beta}_{n,m}(1/2, 1/2)$.

It turns out that the optimal value of (u,v) minimizing the asymptotic variance of $\hat{\beta}_{n,m}(u,v)$ does depend on the unknown parameter β and is moreover not continuous as a function of β. In this paper we instead search for a numerical minimax solution (u_0, v_0) of (u,v) that minimizes

$$\sup_{\beta \in \mathbb{R}} \left\{ \frac{\text{asymptotic variance of } \hat{\beta}_{n,m}(u,v)}{\text{asymptotic variance of } \hat{\beta}_{n,m}^{(P)}} \right\},$$

the maximum asymptotic relative efficiency of $\hat{\beta}_{n,m}^{(P)}$ with respect to $\hat{\beta}_{n,m}(u,v)$.

To reduce the asymptotic variance of $\hat{\beta}_{n,m}(u_0, v_0)$, we then consider convex combinations of form

$$\hat{\beta}_{n,m}(p) := p\hat{\beta}_{n,m}(u_0, v_0) + (1-p)\hat{\beta}_{n,[v_0 m]}(u_0, v_0), \ p \in [0,1]. \tag{1.2}$$

We establish asymptotic normality of $\hat{\beta}_{n,m}(p)$ and determine the optimal choice $p^*(\beta)$ of p minimizing the asymptotic variance of $\hat{\beta}_{n,m}(p)$. The estimators $\hat{\beta}_{n,m}^{(P)}$

and $\hat{\beta}_{n,m}(u_0, v_0)$ are then clearly outperformed by $\hat{\beta}_{n,m}(p^*(\beta))$. Since $p^*(\beta)$ depends on β, a data-driven version of $p^*(\beta)$ is finally plugged into $\hat{\beta}_{n,m}(p^*(\beta))$, with the resulting estimator having the same asymptotic performance as $\hat{\beta}_{n,m}(p^*(\beta))$.

2. Minimax Choice of Generalized Pickands Estimator

We work with an *intermediate sequence* $m = m(n)$ (*i.e.* sequence of integers $m = m(n)$ such that $m \to \infty$ and $m/n \to 0$ as $n \to \infty$). Yun (2002) showed that the generalized Pickands estimator $\hat{\beta}_{n,m}(u, v)$ is weakly consistent for any intermediate sequence $m = m(n)$ and that it is also strongly consistent if the sequence $m = m(n)$ increases suitably rapidly.

Let the function U be defined by $U(x) := F^{-1}(1 - 1/x)$, $x > 1$, where F^{-1} denotes the quantile function of F. Then $F \in \mathcal{D}(G_\beta)$ for some $\beta \in \mathbb{R}$ if and only if there exist functions $a(t) > 0$ and $R(t, x)$ such that, for $x > 0$,

$$\frac{U(tx) - U(t)}{a(t)} = \frac{x^\beta - 1}{\beta} + R(t, x), \; R(t, x) = o(1) \text{ as } t \to \infty \tag{2.1}$$

(*cf.* de Haan, 1984). In this case the function $a(t)$ is regularly varying at infinity with index β [$a(t) \in RV_\beta$]. For asymptotic normality of $\hat{\beta}_{n,m}(u, v)$, one has to consider the second order behavior of U. Among several second order conditions on U introduced in the literature (*cf.* Smith, 1987; Dekkers and de Haan, 1989; Pereira, 1994), the most general form was made by de Haan and Stadtmüller (1996), who assumed that, for $x > 0$,

$$R(t, x) = A(t)H(x) + o(A(t)), \; A(t) = o(1) \text{ as } t \to \infty, \tag{2.2}$$

where $A(t)$ is a function of constant sign for large values of t and

$$H(x) := \frac{1}{\rho}\left(\frac{x^{\beta+\rho} - 1}{\beta + \rho} - \frac{x^\beta - 1}{\beta}\right)$$

for some $\rho \leq 0$. In this case one must have $|A(t)| \in RV_\rho$.

For convenience, we also work with a slightly stronger assumption than $F \in \mathcal{D}(G_\beta)$, *i.e.* we assume that F is differentiable in a left neighborhood of the right endpoint of F and there exist constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} \frac{d}{dx} F^n(a_n x + b_n) = G'_\beta(x)$$

locally uniformly in x with $1 + \beta x > 0$ (*cf.* Pickands, 1986). In this case we say that F belongs to the differentiable domain of attraction of G_β [$F \in \mathcal{D}_{dif}(G_\beta)$].

Clearly, $F \in \mathcal{D}_{dif}(G_\beta)$ implies $F \in \mathcal{D}(G_\beta)$ with the same normalizing constants a_n and b_n.

Under condition (2.2), define $f(t) := t/A^2(t)$ and let $f^{-1}(t)$ denote any asymptotic inverse function of $f(t)$ as $t \to \infty$. For ease of reference, we supply the following result which was established by Yun (2002). By $\xrightarrow{d}$ we denote convergence in distribution.

Theorem 2.1. *Suppose $F \in \mathcal{D}_{dif}(G_\beta)$ for some $\beta \in \mathbb{R}$, so that (2.1) holds. Let $u, v \in (0,1)$.*

(a) *If $R(t,x) \equiv 0$, then*

$$\sqrt{m}(\hat{\beta}_{n,m}(u,v) - \beta) \xrightarrow{d} N(0, \sigma_\beta^2(u,v)) \text{ as } n \to \infty \tag{2.3}$$

for any intermediate sequence $m = m(n)$, where

$$\sigma_\beta^2(u,v) := \frac{(1+u^{-2\beta-1})(1-v) - 2u^{-\beta-1} \cdot \max\{u-v, 0\}}{v \log^2 v} \left(\frac{\beta}{1-u^{-\beta}}\right)^2.$$

(b) *For $R(t,x)$ not being identically zero, assume further that (2.2) holds. Then (2.3) holds for any intermediate sequence $m = m(n)$ such that $m = o(n/f^{-1}(n))$.*

It is worth noting that (2.1) holds with $R(t,x) \equiv 0$ if and only if F is a generalized Pareto distribution function up to a scale and location parameter (*cf.* Theorem 3.1 of Pereira, 1994).

Now the asymptotic variance $\sigma_\beta^2(u,v)$ does not have the same behavior, as a function of (u,v), for all $\beta \in \mathbb{R}$. Thus the optimal value of (u,v) minimizing $\sigma_\beta^2(u,v)$ does depend on the unknown parameter β, which we denote by $(u^*(\beta), v^*(\beta))$. If the function $\beta \mapsto (u^*(\beta), v^*(\beta))$ were continuous, then we could consider plugging data-driven versions of $u^*(\beta)$ and $v^*(\beta)$ into $\hat{\beta}_{n,m}(u^*(\beta), v^*(\beta))$. In general we cannot get the explicit form of $(u^*(\beta), v^*(\beta))$ due to the irregular expression of $\sigma_\beta^2(u,v)$, but it can be numerically seen that the function $\beta \mapsto (u^*(\beta), v^*(\beta))$ is not continuous at $\beta = 0$ by simply computing the values $(u^*(\beta), v^*(\beta))$ around $\beta = 0$. This implies that any adaptive procedure is not feasible in general.

Instead, we search for a numerical minimax solution of (u,v) with respect to asymptotic relative efficiency (ARE). The ARE of $\hat{\beta}_{n,m}^{(P)}$ with respect to $\hat{\beta}_{n,m}(u,v)$

is given by

$$
\begin{aligned}
& g(\beta, u, v) \\
& := \sigma_\beta^2(u,v)/\sigma_\beta^2(1/2,1/2) \\
& = \frac{(1+u^{-2\beta-1})(1-v) - 2u^{-\beta-1}\cdot\max\{u-v,0\}}{2(1+2^{-2\beta-1})v}\left(\frac{1-2^{-\beta}}{1-u^{-\beta}}\right)^2\left(\frac{\log 2}{\log v}\right)^2
\end{aligned}
$$

for $u, v \in (0,1)$. The ARE function $g(\beta, u, v)$ may be interpreted as a loss function with action (u, v). Then a minimax choice of (u, v) is a value of (u, v) that minimizes $\sup_{\beta\in\mathbb{R}} g(\beta, u, v)$, the maximum ARE. Unfortunately, for each value of (u, v) the maximum ARE has no explicit form but can still be computed numerically. In Figure 2.1(a) we plot the maximum ARE function $\sup_{\beta\in\mathbb{R}} g(\beta, u, v)$. It looks somewhat flat in the middle part of the plot, but it is actually convex and has a minimum value 0.7419 (approx.) at $(u, v) = (u_0, v_0) := (1/2, 1/5)$ (approx.) as seen in the enlarged plot of Figure 2.1(b).

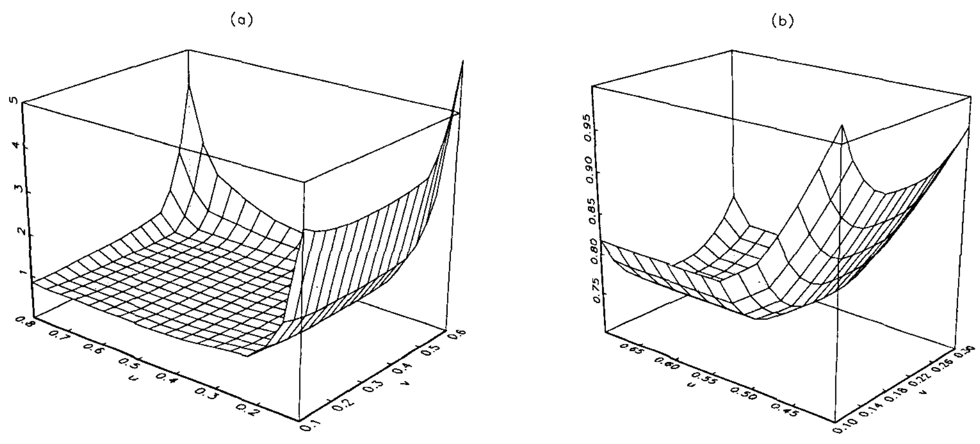

FIGURE 2.1 *Maximum ARE of the Pickands estimator* $\hat\beta_{n,m}^{(P)}$ *with respect to* $\hat\beta_{n,m}(u, v)$

In Figure 2.2 we plot the ARE function $g(\beta, u_0, v_0)$ for the minimax choice (u_0, v_0). It has a minimum value 0.3485 (approx.) at $\beta = -1/2$. Also note that

$$
\lim_{\beta\to\infty} g(\beta, u_0, v_0) = \lim_{\beta\to-\infty} g(\beta, u_0, v_0) = 0.7419 \text{ (approx.)}.
$$

In other words, the asymptotic performance of the minimax estimator $\hat\beta_{n,m}(u_0, v_0)$

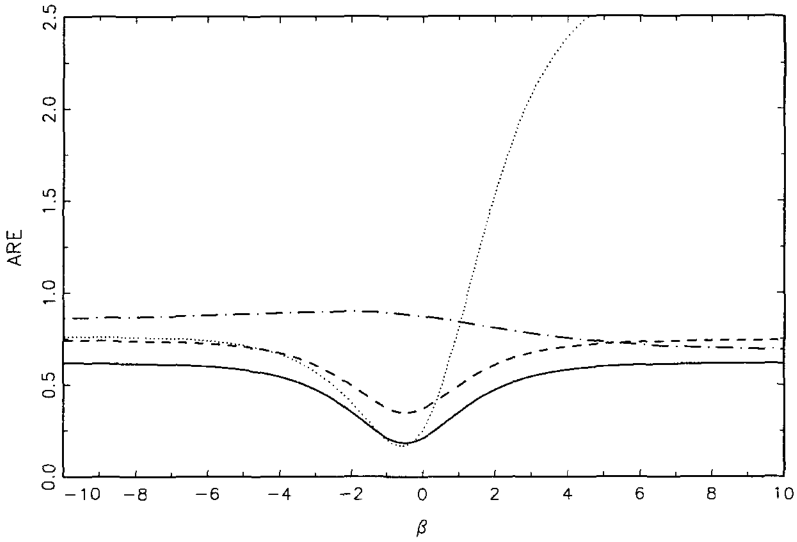

FIGURE 2.2 *ARE of $\hat{\beta}^{(P)}_{n,m}$ with respect to $\hat{\beta}_{n,m}(u_0, v_0)$ (broken line), $\hat{\beta}_{n,m}(\theta_0, \theta_0)$ (dotted line), $\hat{\beta}_{n,m}(1/c_0, c_0/4)$ (dotted and broken line) and $\hat{\beta}_{n,m}(p^*(\tilde{\beta}_n))$ (solid line)*

is better than that of the Pickands estimator $\hat{\beta}^{(P)}_{n,m}$ by at least 25.81% and at most 65.15%. The best improvement occurs when $\beta = -1/2$.

Pereira (1994) and Yun (2000) suggested the use of $\hat{\beta}_{n,m}(\theta_0, \theta_0)$ and $\hat{\beta}_{n,m}(1/c_0, c_0/4)$, respectively, where $\theta_0 := 0.14$ and $c_0 := 1.5984$. Therefore it is natural to compare the asymptotic performance of $\hat{\beta}_{n,m}(u_0, v_0)$ with those of $\hat{\beta}_{n,m}(\theta_0, \theta_0)$ and $\hat{\beta}_{n,m}(1/c_0, c_0/4)$. Figure 2.2 also includes the ARE functions $g(\beta, \theta_0, \theta_0)$ (ARE of $\hat{\beta}^{(P)}_{n,m}$ with respect to $\hat{\beta}_{n,m}(\theta_0, \theta_0)$) and $g(\beta, 1/c_0, c_0/4)$ (ARE of $\hat{\beta}^{(P)}_{n,m}$ with respect to $\hat{\beta}_{n,m}(1/c_0, c_0/4)$).

The plot exhibits that the Pereira estimator $\hat{\beta}_{n,m}(\theta_0, \theta_0)$ outperforms the minimax estimator $\hat{\beta}_{n,m}(u_0, v_0)$ only for $\beta \in (-4.3631, 0.3585)$ and is worse than even the Pickands estimator $\hat{\beta}^{(P)}_{n,m}$ for $\beta > 1.24$. In fact, we have

$$\lim_{\beta \to \infty} g(\beta, \theta_0, \theta_0) = 2.7268 \text{ (approx.)},$$

$$\lim_{\beta \to -\infty} g(\beta, \theta_0, \theta_0) = 0.7635 \text{ (approx.)},$$

$$\lim_{\beta\to\infty} g(\beta, 1/c_0, c_0/4) = 0.6857 \text{ (approx.)},$$

$$\lim_{\beta\to-\infty} g(\beta, 1/c_0, c_0/4) = 0.8579 \text{ (approx.)}.$$

Therefore it is clear that, among the estimators compared so far, $\hat{\beta}_{n,m}(u_0, v_0)$ is the only estimator that gives a fairly good asymptotic performance for all real values of β. These considerations may support a general preference of the minimax estimator $\hat{\beta}_{n,m}(u_0, v_0)$ particularly for conservative analysts.

3. Convex Combinations of Minimax Estimator

Though the minimax estimator $\hat{\beta}_{n,m}(u_0, v_0)$ shows a fairly good performance, its asymptotic efficiency may be further improved by convex combinations. In this section we consider simple convex combinations given in (1.2).

The following lemma is needed to prove asymptotic normality of $\hat{\beta}_{n,m}(p)$.

Lemma 3.1. (Cooil, 1985) *If $F \in \mathcal{D}_{dif}(G_\beta)$ for some $\beta \in \mathbb{R}$, then for any intermediate sequence $m = m(n)$, the stochastic process $\{Z_{n,m}(\theta), \theta \in (0,1]\}$ defined by*

$$Z_{n,m}(\theta) := \frac{\sqrt{m}(X^{(n)}_{[\theta m]} - U(n/(\theta m)))}{a(n/m)}$$

with $a(\cdot)$ in (2.1) converges to the Gaussian process $\{Z(\theta), \theta \in (0,1]\}$ defined by $Z(\theta) := \theta^{-\beta-1}W(\theta)$, where $\{W(\theta), \theta \geq 0\}$ is a standard Wiener process, in the sense that for every $k = 1, 2, ...$ and any values $0 < \theta_1 < \cdots < \theta_k \leq 1$,

$$(Z_{n,m}(\theta_1), ..., Z_{n,m}(\theta_k)) \xrightarrow{d} (Z(\theta_1), ..., Z(\theta_k)) \text{ as } n \to \infty.$$

Theorem 3.1. *Suppose $F \in \mathcal{D}_{dif}(G_\beta)$ for some $\beta \in \mathbb{R}$, so that (2.1) holds. Let $p \in [0,1]$.*

(a) *If $R(t,x) \equiv 0$, then*

$$\sqrt{m}(\hat{\beta}_{n,m}(p) - \beta) \xrightarrow{d} N(0, \sigma^2_\beta(u_0, v_0)\nu^2_\beta(p)) \text{ as } n \to \infty \tag{3.1}$$

for any intermediate sequence $m = m(n)$, where

$$\sigma^2_\beta(u_0, v_0) = \frac{4 - 3\cdot 2^{\beta+1} + 2^{2\beta+3}}{\log^2 5}\left(\frac{\beta}{1-2^\beta}\right)^2,$$

$$\nu^2_\beta(p) := 1 - (1-p)\frac{2 - 2^{\beta-1} + 2^{2\beta+2}}{1 - 3\cdot 2^{\beta-1} + 2^{2\beta+1}} + (1-p)^2\frac{6 - 13\cdot 2^{\beta-1} + 3\cdot 2^{2\beta+2}}{1 - 3\cdot 2^{\beta-1} + 2^{2\beta+1}}.$$

(b) *For $R(t,x)$ not being identically zero, assume further that (2.2) holds. Then (3.1) holds for any intermediate sequence $m = m(n)$ such that $m = o(n/f^{-1}(n))$.*

Proof. Let $m = m(n)$ be any intermediate sequence and write

$$A_{n,m} := \frac{X_m^{(n)} - X_{[u_0 m]}^{(n)}}{X_{[v_0 m]}^{(n)} - X_{[u_0 v_0 m]}^{(n)}}.$$

We will see below that $A_{n,m} - v_0^\beta$ and $A_{n,[v_0 m]} - v_0^\beta$ are of order $O_p(m^{-1/2})$, and so by applying the expansion $\log(1+x) = x + o(x)$ as $x \to 0$ we have

$$\begin{aligned}
&\sqrt{m}(\hat{\beta}_{n,m}(p) - \beta) \\
&= \frac{\sqrt{m}}{\log v_0}\left\{p \log\left(1 + \frac{A_{n,m} - v_0^\beta}{v_0^\beta}\right) + (1-p)\log\left(1 + \frac{A_{n,[v_0 m]} - v_0^\beta}{v_0^\beta}\right)\right\} \\
&= \frac{1}{v_0^\beta \log v_0}\left\{p\sqrt{m}(A_{n,m} - v_0^\beta) + (1-p)\sqrt{m}(A_{n,[v_0 m]} - v_0^\beta)\right\} + o_p(1)
\end{aligned}$$

as $n \to \infty$. Since (2.1) and Lemma 3.1 lead to

$$\begin{aligned}
&\frac{X_{[v_0 m]}^{(n)} - X_{[u_0 v_0 m]}^{(n)}}{a(n/m)} \\
&= \frac{U(n/(v_0 m)) - U(n/(u_0 v_0 m))}{a(n/m)} + \frac{1}{\sqrt{m}}(Z_{n,m}(v_0) - Z_{n,m}(u_0 v_0)) \\
&= \frac{v_0^{-\beta}(1 - u_0^{-\beta})}{\beta} + o_p(1) \text{ as } n \to \infty,
\end{aligned}$$

it follows that, as $n \to \infty$,

$$A_{n,m} - v_0^\beta \overset{p}{\sim} \frac{\beta}{v_0^{-\beta}(1 - u_0^{-\beta})} \times \frac{X_m^{(n)} - X_{[u_0 m]}^{(n)} - v_0^\beta(X_{[v_0 m]}^{(n)} - X_{[u_0 v_0 m]}^{(n)})}{a(n/m)}.$$

Similarly, one can show that, as $n \to \infty$,

$$A_{n,[v_0 m]} - v_0^\beta \overset{p}{\sim} \frac{\beta}{v_0^{-2\beta}(1 - u_0^{-\beta})} \times \frac{X_{[v_0 m]}^{(n)} - X_{[u_0 v_0 m]}^{(n)} - v_0^\beta(X_{[v_0^2 m]}^{(n)} - X_{[u_0 v_0^2 m]}^{(n)})}{a(n/m)}.$$

Thus,

$$
\begin{aligned}
&\sqrt{m}(\hat{\beta}_{n,m}(p) - \beta) \\
&\overset{p}{\sim} \frac{\beta}{(1-u_0^{-\beta})\log v_0} \times \frac{\sqrt{m}}{a(n/m)} \Big\{p(X_m^{(n)} - X_{[u_0 m]}^{(n)}) \\
&\quad + (1-2p)v_0^{\beta}(X_{[v_0 m]}^{(n)} - X_{[u_0 v_0 m]}^{(n)}) - (1-p)v_0^{2\beta}(X_{[v_0^2 m]}^{(n)} - X_{[u_0 v_0^2 m]}^{(n)})\Big\} + o_p(1) \\
&= \frac{\beta}{(1-u_0^{-\beta})\log v_0}(B_{n,m} + C_{n,m}) + o_p(1) \text{ as } n \to \infty, \qquad (3.2)
\end{aligned}
$$

where

$$
\begin{aligned}
B_{n,m} &= p(Z_{n,m}(1) - Z_{n,m}(u_0)) + (1-2p)v_0^{\beta}\left\{Z_{n,m}(v_0) - Z_{n,m}(u_0 v_0)\right\} \\
&\quad -(1-p)v_0^{2\beta}\left\{Z_{n,m}(v_0^2) - Z_{n,m}(u_0 v_0^2)\right\}, \\
C_{n,m} &= \frac{\sqrt{m}}{a(n/m)}\Big[p\left\{U(n/m) - U(n/(u_0 m))\right\} \\
&\quad + (1-2p)v_0^{\beta}\left\{U(n/(v_0 m)) - U(n/(u_0 v_0 m))\right\} \\
&\quad - (1-p)v_0^{2\beta}\left\{U(n/(v_0^2 m)) - U(n/(u_0 v_0^2 m))\right\}\Big].
\end{aligned}
$$

From Lemma 3.1, it is readily checked that

$$
B_{n,m} \overset{d}{\to} N(0, V) \text{ as } n \to \infty, \qquad (3.3)
$$

where

$$
V = (4 - 3\cdot 2^{\beta+1} + 2^{2\beta+3}) - (1-p)(8 - 2^{\beta+1} + 2^{2\beta+4}) + (1-p)^2(24 - 13\cdot 2^{\beta+1} + 3\cdot 2^{2\beta+4}).
$$

(a) The assumption that $R(t,x) \equiv 0$ implies $C_{n,m} \equiv 0$ by (2.1). Hence, from (3.2) and (3.3), we have, as $n \to \infty$,

$$
\sqrt{m}(\hat{\beta}_{n,m}(p) - \beta) \overset{d}{\to} N(0, V\beta^2/((1-u_0^{-\beta})^2 \log^2 v_0)) = N(0, \sigma_\beta^2(u_0, v_0)\nu_\beta^2(p)).
$$

(b) Note that $m = o(n/f^{-1}(n))$ if and only if $\lim_{n\to\infty} \sqrt{m}A(n/m) = 0$ (*cf.* Lemma 2.1 of Yun, 2002). From (2.1) and (2.2), we thus have

$$
\begin{aligned}
C_{n,m} &= \Big[(1-2p)v_0^{\beta}\left\{H(v_0^{-1}) - H(u_0^{-1}v_0^{-1})\right\} - pH(u_0^{-1}) \\
&\quad - (1-p)v_0^{2\beta}\left\{H(v_0^{-2}) - H(u_0^{-1}v_0^{-2})\right\}\Big]\sqrt{m}A(n/m) + \sqrt{m}o(A(n/m)) \\
&\to 0 \text{ as } n \to \infty.
\end{aligned}
$$

Hence, (3.1) follows from (3.2) and (3.3) as in case (a). □

The factor $\nu_\beta^2(p)$ is now the ARE of $\hat\beta_{n,m}(u_0, v_0)$ with respect to $\hat\beta_{n,m}(p)$. The optimal choice of p minimizing $\nu_\beta^2(p)$ is

$$p^*(\beta) := 1 - \frac{1 - 2^{\beta-2} + 2^{2\beta+1}}{6 - 13 \cdot 2^{\beta-1} + 3 \cdot 2^{2\beta+2}},$$

in which case $\nu_\beta^2(p)$ becomes

$$\nu_\beta^2(p^*(\beta)) = 1 - (1 - p^*(\beta)) \frac{1 - 2^{\beta-2} + 2^{2\beta+1}}{1 - 3 \cdot 2^{\beta-1} + 2^{2\beta+1}}.$$

Since $\nu_\beta^2(p^*(\beta)) < 1$ for all $\beta \in \mathbb{R}$, the convex combination $\hat\beta_{n,m}(p^*(\beta))$ is clearly superior to the minimax estimator $\hat\beta_{n,m}(u_0, v_0)$.

In Figure 3.1 we display the ARE function $\nu_\beta^2(p^*(\beta))$ as well as the optimal weight $p^*(\beta)$. Notice that $\nu_\beta^2(p^*(\beta))$ has a minimum value 0.522 (approx.) at $\beta = -1/2$ and that

$$\lim_{\beta\to\infty} \nu_\beta^2(p^*(\beta)) = \lim_{\beta\to-\infty} \nu_\beta^2(p^*(\beta)) = 5/6.$$

The weights $p^*(\beta)$ range between 0.7538 (approx.) and 5/6.

Since the optimal weight depends on the unknown β, it is reasonable to utilize the adaptive estimator $p^*(\tilde\beta_n)$, where $\tilde\beta_n$ is an initial estimator of β which is weakly consistent. In practice, one may use as an initial estimator

$$\tilde\beta_n = \hat\beta_{n,m}(p^*(0)) = \hat\beta_{n,m}(35/46).$$

The following result implies that the adaptive estimator $\hat\beta_{n,m}(p^*(\tilde\beta_n))$ has the same asymptotic performance as the convex combination $\hat\beta_{n,m}(p^*(\beta))$ with underlying β. For comparison with the estimators discussed in Section 2, a plot of the ARE of $\hat\beta_{n,m}^{(P)}$ with respect to $\hat\beta_{n,m}(p^*(\tilde\beta_n))$, which is given by $g(\beta, u_0, v_0)\nu_\beta^2(p^*(\beta))$, is included in Figure 2.2.

Theorem 3.2. *Assume that the conditions of Theorem 3.1 hold for some $\beta \in \mathbb{R}$. Let $\tilde\beta_n$ be any initial estimator of β which is weakly consistent. Then*

$$\sqrt{m}(\hat\beta_{n,m}(p^*(\tilde\beta_n)) - \beta) \xrightarrow{d} N(0, \sigma_\beta^2(u_0, v_0)\nu_\beta^2(p^*(\beta))) \text{ as } n \to \infty$$

for any intermediate sequence $m = m(n)$ in case (a) and any intermediate sequence $m = m(n)$ such that $m = o(n/f^{-1}(n))$ in case (b).

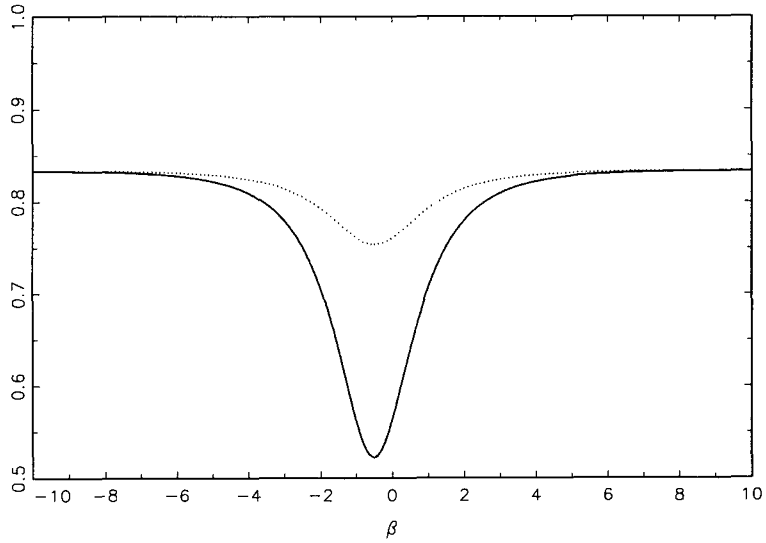

FIGURE 3.1 *ARE of $\hat{\beta}_{n,m}(u_0, v_0)$ with respect to $\hat{\beta}_{n,m}(p^*(\beta))$ (solid line) and optimal weight $p^*(\beta)$ (dotted line)*

Proof. By $\xrightarrow{p}$ we denote convergence in probability. Since $p^*(\cdot)$ is a continuous function, $p^*(\tilde{\beta}_n) \xrightarrow{p} p^*(\beta)$ as $n \to \infty$. It thus follows from Theorem 2.1 that

$$\begin{aligned}
&\sqrt{m}(\hat{\beta}_{n,m}(p^*(\tilde{\beta}_n)) - \hat{\beta}_{n,m}(p^*(\beta))) \\
&= (p^*(\tilde{\beta}_n) - p^*(\beta))\sqrt{m}(\hat{\beta}_{n,m}(u_0, v_0) - \beta) \\
&\quad -(p^*(\tilde{\beta}_n) - p^*(\beta))\sqrt{m}(\hat{\beta}_{n,[v_0 m]}(u_0, v_0) - \beta) \\
&= o_p(1) \text{ as } n \to \infty.
\end{aligned}$$

Hence the assertion is immediate from Theorem 3.1. □

4. Discussions

In the present paper, a minimax criterion in terms of ARE was first used to select a fixed value (u_0, v_0) and then simple convex combinations $\hat{\beta}_{n,m}(p)$ were considered to reduce the asymptotic variance of $\hat{\beta}_{n,m}(u_0, v_0)$. A natural extension to utilize the remaining sample data is to consider higher-order convex combinations of form

$$\hat{\beta}_{n,m}(p_1, ..., p_k) = \sum_{i=1}^{k} p_i \hat{\beta}_{n,[v_0^{i-1}m]}(u_0, v_0), \; p_i \in [0,1], \; \sum_{i=1}^{k} p_i = 1,$$

of length k. Since

$$\hat{\beta}_{n,m}(p_1, ..., p_k) = \sum_{i=1}^{k-1} i(p_i - p_{i+1})\hat{\beta}_{n,m}(u_0, v_0^i) + kp_k\hat{\beta}_{n,m}(u_0, v_0^k),$$

a further generalization, as the referee suggested, is to consider

$$\hat{\beta}_{n,m}(\pi) = \int\int \hat{\beta}_{n,m}(u,v)\,\pi(du, dv),$$

where π is a (possibly signed) measure on the Borel σ-field in $(0,1)^2$ such that $\pi((0,1)^2) = 1$, which will clearly have a smaller variance than $\hat{\beta}_{n,m}(u_0, v_0)$ if the measure π is chosen appropriately. It seems difficult to determine an optimal π minimizing the asymptotic variance of $\hat{\beta}_{n,m}(\pi)$, which will certainly depend on the unknown β, and will be an interesting topic for future research.

Another issue to be worthy of note is how to choose m in practice. For simplicity, the present paper dealt with only the estimators of β which are asymptotically unbiased. Increasing m to obtain a smaller variance of the estimator $\hat{\beta}_{n,m}(p)$ however causes the estimator to have a bias different from zero. In fact, it is possible to determine a theoretically optimal value of m which minimizes the asymptotic mean squared error of $\hat{\beta}_{n,m}(p)$ (*cf.* Yun, 2002). The asymptotic bias of $\hat{\beta}_{n,m}(p)$ turns out to depend on the second order parameter ρ in (2.2) as well as β and p, and so does the optimal value of m as well. Thus, determining the best estimator minimizing the asymptotic mean squared error with the optimal value of m involves a difficult problem of estimating β and ρ simultaneously, which is beyond the scope of the present paper. At this stage, we can say at the least that the higher-order convex combination $\hat{\beta}_{n,m}(p_1, ..., p_k)$ is more robust against an inappropriate choice of m as k increases larger.

Acknowledgements

I am grateful to the referee for his valuable comments.

References

Alves, M. I. F. (1995). "Estimation of the tail parameter in the domain of attraction of an extremal distribution", *Journal of Statistical Planning and Inference*, **45**, 143–173.

Cooil, B. (1985). "Limiting multivariate distributions of intermediate order statistics", *The Annals of Probability*, **13**, 469–477.

de Haan, L. (1984). "Slow variation and characterization of domains of attraction", In *Statistical Extremes and Applications* (J. Tiago de Oliveira, ed.), 31–48, Reidel, Dordrecht.

de Haan, L. and Stadtmüller, U. (1996). "Generalized regular variation of second order", *Journal of the Australian Mathematical Society*, **A61**, 381–395.

Dekkers, A. L. M. and de Haan, L. (1989). "On the estimation of the extreme-value index and large quantile estimation", *The Annals of Statistics*, **17**, 1795–1832.

Drees, H. (1995). "Refined Pickands estimators of the extreme value index", *The Annals of Statistics*, **23**, 2059–2080.

Falk, M. (1994). "Efficiency of convex combinations of Pickands estimator of the extreme value index", *Journal of Nonparametric Statistics*, **4**, 133–147.

Hill, B. M. (1975). "A simple general approach to inference about the tail of a distribution", *The Annals of Statistics*, **3**, 1163–1174.

Pereira, T. T. (1994). "Second order behavior of domains of attraction and the bias of generalized Pickands' estimator", In *Extreme Value Theory and Applications III* (J. Galambos, J. Lechner and E. Simiu, eds.), 165–177, NIST special publication 866.

Pickands, J. (1975). "Statistical inference using extreme order statistics", *The Annals of Statistics*, **3**, 119–131.

Pickands, J. (1986). "The continuous and differentiable domains of attraction of the extreme value distributions", *The Annals of Probability*, **14**, 996–1004.

Smith, R. L. (1987). "Estimating tails of probability distributions", *The Annals of Statistics*, **15**, 1174–1207.

Yun, S. (2000). "A class of Pickands-type estimators for the extreme value index", *Journal of Statistical Planning and Inference*, **83**, 113–124.

Yun, S. (2002). "On a generalized Pickands estimator of the extreme value index", *Journal of Statistical Planning and Inference*, **102**, 389–409.